

TES DIAGNOSTIK (*DIAGNOSTIC TEST*)

Oleh: Risanto Siswosudarmo
Departemen Obstetrika dan Ginekologi
FK UGM Yogyakarta

Pendahuluan.

Test diagnostik adalah sebuah cara (alat) untuk menentukan apakah seseorang menderita penyakit atau tidak, berdasar adanya tanda dan gejala pada orang tersebut. *Test* skrining adalah sebuah cara untuk mengetahui atau mengidentifikasi apakah seseorang yang masih asimtomatik menderita suatu penyakit atau tidak. Tanpa skrining, diagnosis suatu penyakit hanya bisa ditegakkan setelah muncul tanda dan gejala, padahal sebuah penyakit telah ada jauh sebelum tanda dan gejala muncul yang sebenarnya dapat diketahui kalau kita melakukan skrining. Waktu antara kemungkinan terdeteksi secara awal lewat skrining dan deteksi kemudian setelah munculnya tanda dan gejala disebut "*detectable pre-clinical phase*" atau **DPCP**. Jika sebuah penyakit dapat diketahui pada masa DPCP maka *treatment* bisa dilakukan lebih awal dan *outcome* nya pun lebih baik.(1) *Test* diagnostik dilakukan setelah seseorang dinyatakan positif pada *test* skrining untuk menegakkan diagnosis secara lebih pasti (definitif).

Beberapa contoh *test* skrining adalah:

- a. *Pap's smear* dan IVA untuk kanker serviks.
- b. Pemeriksaan gula darah puasa untuk diabetes.
- c. Tekanan okular untuk glaukoma.
- d. Darah dalam feses untuk kanker kolon.
- e. Mamografi untuk kanker payudara.

Kapan sebuah tes skrining bermanfaat?

1. Penyakitnya serius (misal kanker serviks, hipertensi dll).
2. Bila *treatment* dilakukan sebelum tanda dan gejala muncul hasilnya lebih efektif dibanding bila dilakukan setelah tanda dan gejala muncul.
3. Bila prevalensi penyakit dalam DPCP relatif tinggi.

Syarat sebuah *test* skrining yang baik

- Murah

- Mudah dilakukan
- Hanya sedikit memberi rasa tidak nyaman
- *Valid* (bisa membedakan sakit dengan tidak sakit)
- *Reliable* (konsisten, istikomah)

Beberapa istilah

1. **Validitas:** Adalah kemampuan tes untuk menunjukkan dengan benar (akurat) individu mana yang menderita sakit dan mana yang tidak. Validitas tes dicerminkan dengan sensitivitas dan spesifisitas
2. **Sensitivitas:** Adalah kemampuan tes untuk menunjukkan individu mana yang menderita sakit dari seluruh populasi yang benar-benar sakit.
3. **Spesifisitas:** Adalah kemampuan tes untuk menunjukkan individu mana yang **tidak** menderita sakit dari mereka yang benar-benar tidak sakit.
4. **Gold standard test:** Adalah tes terbaik yang tersedia, diterima secara luas yang umumnya kurang nyaman, mahal dan invasif.

Sebuah tes skrining yang ideal adalah yang mempunyai sensitivitas dan spesifisitas tinggi yang berarti validitasnya juga tinggi. Validitas sebuah tes skrining didasarkan atas akurasi dalam mengidentifikasi individu ke dalam sakit dan tidak sakit. Untuk tujuan ini sebuah tes skrining harus dibandingkan dengan sebuah atau beberapa *gold standard test* yang menyatakan bahwa seseorang adalah benar-benar sakit atau tidak sakit. Sayangnya *gold standard test* adalah sebuah alat diagnostik yang sering kali kurang nyaman, mahal dan invasif.

Perhitungan dengan menggunakan tabel 2X2.

Tes baru (tes diagnostik)	<i>Gold standard test</i>		Total
	Sakit	Tidak sakit	
Positif	a	b	a+b
Negatif	c	d	c+d
Total	a+c	b+d	N

Sensitivitas: $a/(a+c)$

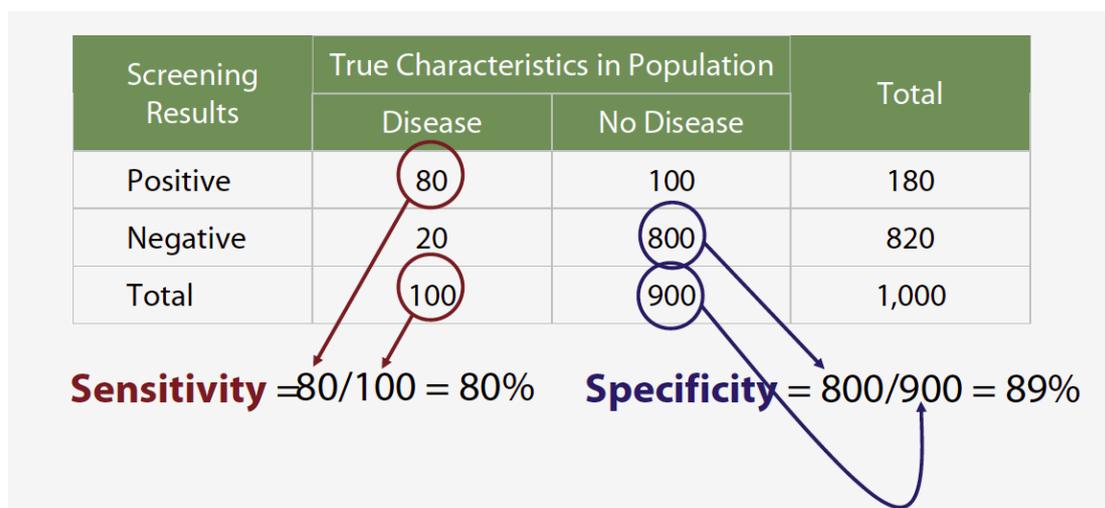
Spesifisitas: $d/(b+d)$

Contoh:

- Diasumsikan dalam 1000 populasi, terdapat 100 orang dengan penyakit dan 900 tanpa penyakit (prevalensi 10%).
- Sebuah tes skrining dipakai untuk mengidentifikasi 100 orang yang sakit.
- Hasilnya adalah sebagai berikut:

Hasil skrining	Keadaan sebenarnya dalam populasi		Total
	Sakit	Tidak sakit	
Positif	80	100	180
Negatif	20	800	820
Total	100	900	1000

Hasilnya sebagai berikut:



5. **Positive predictive value** (PPV) atau nilai ramal positif (NRP). Adalah proporsi pasien yang tes nya positif dan betul menderita sakit. Dengan kata lain “Jika tes seseorang positif, berapa probabilitas dia betul-betul menderita penyakit?” Rumus: $PPV = a/(a+b)$.
6. **Negative predictive value** (NPV) atau nilai ramal negatif (NRN). Adalah proporsi pasien yang tes nya negatif dan betul-betul tidak menderita sakit. Bisa juga dikatakan “Jika tes seseorang negatif, berapa probabilitas dia betul-betul tidak menderita penyakit?” Rumus: $NPV = d/(c+d)$.

7. **SnNOUT**. Jika sebuah tes dengan **sensitivitas** tinggi adalah **negatif**, maka kemungkinan untuk me **rule out** sebuah penyakit adalah besar (tes ini efektif untuk mengatakan bahwa sangat mungkin pasien benar-benar tidak menderita sakit).
8. **SpPIN**. Jika sebuah tes dengan spesifisitas tinggi adalah positif, maka kemungkinan untuk me **rule in** sebuah penyakit adalah besar (tes ini sangat efektif untuk mengatakan bahwa sangat mungkin pasien benar-benar menderita sakit).
9. **True positive**: Adalah mereka yang tes nya positif dan berpenyakit. *True positive rate* (TPR) adalah proporsi mereka yang tes nya positif terhadap seluruh populasi yang berpenyakit. Rumus: **TPR = a/(a+c)**. **TPR tidak lain adalah sensitivitas**.
10. **False positive**: Adalah mereka yang tes nya positif padahal sebenarnya mereka tidak berpenyakit. *False positive rate* (FPR) adalah proporsi mereka yang tes nya positif terhadap seluruh populasi yang tidak berpenyakit.
Rumusnya FPR = b/(b+d). Ternyata **FPR = 1 – spesifisitas**.
Catatan: TPR dan FPR dipakai untuk menghitung receiver operating characteristic curve atau ROC curve. Lihat bawah.
11. **True negative**: Adalah mereka yang tesnya negatif dan benar-benar tidak berpenyakit. *True negative rate* (TNR) adalah proporsi mereka yang tesnya negatif terhadap populasi yang benar-benar tidak sakit. Rumusnya **TNR = d/(b+d)**, dan ini adalah **spesifisitas**.
12. **False negative**: Adalah mereka yang *test* nya negatif padahal sebenarnya mereka berpenyakit. *False negative rate* (FNR) adalah proporsi mereka yang tesnya negatif terhadap seluruh populasi yang berpenyakit.
Rumusnya **FNR = d/(b+d)**.
13. **Likelihood ratio (LR)**.
 - a. **LR (+)** adalah **rasio** antara probabilitas tes yang positif pada individu yang berpenyakit dengan probabilitas tes yang positif pada individu yang tidak berpenyakit. Dengan kata lain LR untuk hasil test (+) menunjukkan **berapa kali kemungkinan hasil tes (+) terjadi pada kelompok populasi yang berpenyakit dibanding dengan hasil tes (+) pada kelompok populasi yang tidak berpenyakit**. Bisa juga dikatakan **LR (+) adalah rasio antara true positive rate (TPR) dengan false positive rate (FPR)**. **LR (+) harus lebih**

besar dari 1, dan tes yang baik harus mempunyai LR (+) yang besar, sebab hasil tes (+) pada kelompok populasi yang berpenyakit harus lebih besar dibanding pada kelompok yang tidak berpenyakit.

Rumusnya: *Likelihood ratio positive* (LR+) = Sensitivity/(1-Specificity).

LR (+) adalah indikator yang terbaik untuk memasukkan kemungkinan seseorang menderita penyakit (*ruling in*). Makin besar LR (+) makin, makin besar kemungkinan seseorang menderita penyakit (kemungkinan diagnosisnya betul makin besar). Tes diagnostik yang baik adalah TD dengan LR (+) > 10, sehingga dengan hasil tes positif kemungkinan diagnosis betul makin besar.

- b. LR (-) adalah **rasio** antara probabilitas hasil tes negatif pada individu yang berpenyakit dengan probabilitas hasil tes negatif pada individu yang tidak berpenyakit. Dengan kata lain LR (-) menunjukkan berapa kali **lebih jarang** sebuah hasil tes (-) terjadi pada kelompok yang berpenyakit dibanding kelompok yang tidak berpenyakit. Bisa juga dikatakan LR (-) adalah **rasio antara *false negative rate* (FNR) pada individu yang berpenyakit dengan *false negative rate* (FNR) pada individu yang tidak berpenyakit**. LR (-) biasanya kurang dari 1, dan makin kecil LR (-) maka tes nya makin baik. Mengapa? Karena hasil tes negatif seharusnya lebih jarang terjadi pada kelompok yang berpenyakit dibanding dengan kelompok yang tidak berpenyakit.

Rumusnya: *Likelihood ratio negative* (LR-) = (1-Sensitivity)/Specificity

Berapa nilai LR yang ideal? Perhatikan tabel di bawah ini.

LR(+)	LR(-)	Impact on likelihood
10	0.1	Excellent
6	0.2	Very good
2	0.5	Fair
1	1	Useless

Sensitivitas dan spesifisitas dipengaruhi oleh *cut off point* untuk menentukan kriteria positivities, yakni nilai berapa seseorang dikatakan sakit atau berpenyakit. Sebaliknya **nilai ramal positif**, sangat tergantung pada **prevalensi penyakit yang ada dalam populasi** yang sedang dites. Makin tinggi prevalensi penyakit dalam

populasi, makin tinggi pula PPV. Dengan demikian cara utama untuk menaikkan hasil dalam sebuah skrining adalah dengan menargetkan tes pada kelompok orang yang berisiko tinggi menderita penyakit.

Berbicara tentang nilai ramal, perhatikan gambar berikut:

When thinking about
Predictive Value of a Test, ...

... imagine you are a physician discussing the results of a screening test with a patient.



+ 1) If the test was *positive*, how likely is it that he really *has* the disease? [How *worried* should he be?]



- 2) If the test was *negative*, how likely is it that he really *does NOT* have it? [How *reassured* should he be?]



Lihat skema berikut ini:

		Disease	
		+	-
Test	+	a (True positives)	b (False positives)
	-	c (False negatives)	d (True negatives)

Dengan contoh di atas, perhitungan PPV dan NPV adalah sebagai berikut:

Screening Results	True Characteristics in Population		Total
	Disease	No Disease	
Positive	80	100	180
Negative	20	800	820
Total	100	900	1,000

$$PPV = a/(a+b) = 80/180 = 44,44\%$$

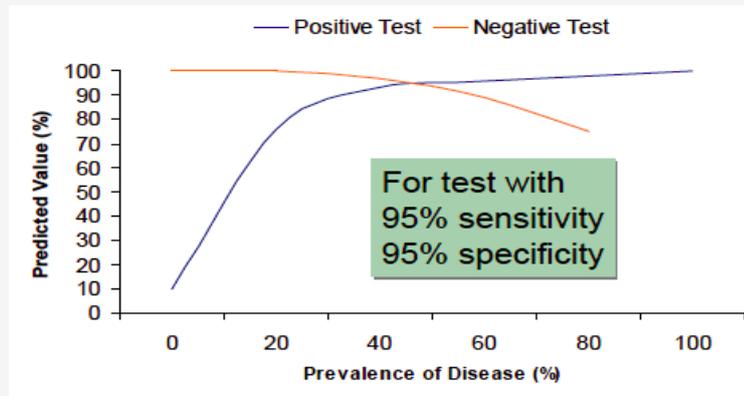
$$NPV = d/(c+d) = 20/820 = 2,44\%$$

Predictive value dipengaruhi oleh prevalensi penyakit, sebagaimana terlihat dalam tabel berikut:

Relationship of Disease Prevalence to Predictive Value

Example: Sensitivity = 99%; Specificity = 95%					
Disease Prevalence	Test Results	Sick	Not Sick	Totals	Positive Predictive Value
1%	+	99	495	594	$\frac{99}{594} = 17\%$
	-	1	9,405	9,406	
	Totals	100	9,900	10,000	
5%	+	495	475	970	$\frac{495}{970} = 51\%$
	-	5	9,025	9,303	
	Totals	500	9,500	10,000	

Prevalence of Disease



Adapted from Mausner JS, Kramer S. Epidemiology: an Introductory Text. Philadelphia, WB Saunders 1985, p221.

Jadi, jika *test* seseorang positif:

- Probabilitas dia menderita penyakit (PPV) tergantung pada prevalensi penyakit pada populasi di mana tes dilakukan, tetapi juga pada validitas tes (yakni sensitivitas dan spesifisitas).
- Secara umum spesifisitas mempunyai pengaruh lebih besar terhadap nilai prediktif.

Contoh lain perubahan PPV berdasar prevalensi penyakit.

Skenario I.

Skrining pada donor darah perempuan yang prevalensi HIV (+) adalah 0,01% (hanya 10 HIV positif dari 100.000 donor perempuan). *Test* skrining memiliki sensitivitas 100% dan spesifisitas 95%. Berapa PPV nya?

Perhatikan tabel berikut:

Tes skrining	<i>Gold Standard</i>		Total Kolom
	HIV (+)	HIV (-)	
Tes (+)	10	510	520
Tes (-)	0	99.480	99.480
Total Lajur	10	99.990	100.000

Jawab: PPV = 10/520 = 1,9%

Skenario II.

Skrining dilakukan pada laki-laki yang datang di klinik penyakit menular seksual yang prevalensi HIV (+) adalah 4 %. Tes skrining memiliki sensitivitas dan spesifisitas yang sama yakni masing-masing 100% dan 95%. Berapa PPV nya?

Perhatikan tabel berikut ini:

Tes skrining	<i>Gold Standard</i>		Total Kolom
	HIV (+)	HIV (-)	
Tes (+)	4.000	480	4.480
Tes (-)	0	95.520	95.520
Total Lajur	4.000	96.000	100.000

Jawab: PPV = 4000/4480 = 89%.

Skenario III.

Skrining dilakukan pada pengguna obat suntik intravena yang prevalensi HIV (+) nya adalah 20%. Berapa PPV nya?

Perhatikan tabel di bawah ini:

Tes skrining	<i>Gold Standard</i>		Total Kolom
	HIV (+)	HIV (-)	
Tes (+)	20.000	400	20.400
Tes (-)	0	79.600	95.520
Total Lajur	20.000	80.000	100.000

Jawab: PPV = 20.000/20.400 = 98.%

Apa arti ketiga skenario di atas? Jawabnya adalah: Jika anda mempunyai sumber daya yang terbatas untuk melakukan skrining dan anda ingin mendapatkan hasil yang maksimal maka anda harus menyasar kelompok populasi yang prevalensi penyakitnya tinggi, bukan pada kelompok yang kemungkinan timbulnya penyakit (prevalensinya) rendah.

Efek samping sebuah skrining.

Sepintas kelihatan dengan skrining kita bisa mengetahui berapa kemungkinan seseorang menderita sakit, sehingga ini merupakan hal yang harus selalu dikerjakan.

Meskipun demikian skrining membawa konsekuensi biaya, sehingga harus ditimbang antara manfaat dan mudaratnya, terutama pada subkelompok yang prevalensi penyakitnya rendah.

Ada dua hal yang harus diperhatikan yakni: *false positive* (positif palsu) dan *false negative* (negatif palsu).

- *False positives*, yakni mereka yang tesnya positif tetapi sebenarnya tidak berpenyakit.
- *False negatives*, yakni mereka yang tesnya negatif padahal sebenarnya mereka berpenyakit.

Konsekuensi dari *false positive* adalah pasien merasa takut yang berlebihan sampai kepada cara diagnosis yang mahal dan invasif atau bahkan pengobatan yang tidak perlu (*over treatment*). Sebaliknya bila *test* nya adalah *false negative*, pasien akan merasa aman-aman saja, padahal ia sebenarnya mempunyai penyakit. Akibatnya diagnosis menjadi terlambat dan pengobatan yang seharusnya dilakukan menjadi tidak dilakukan sehingga morbiditas dan mortalitas pun dan akhirnya biaya yang dibutuhkan juga meningkat.

Receiver operating characteristic curve (ROC) curve.

ROC curve menunjukkan beberapa hal:

1. Titik temu tawar menawar (*trade off*) antara **true positive rate** (sensitivitas) dengan **false positive rate** (1-spesifisitas), karena setiap peningkatan sensitivitas akan disertai dengan penurunan spesifisitas. Dalam hal ini sumbu **Y adalah sensitivitas (true positif rate)** dan sumbu **X nya adalah 1-spesifisitas (false positive rate)**.
2. Makin kurva mendekati ke kiri dan ke atas ruang ROC, tes makin akurat.
3. Makin kurva mendekati garis diagonal tes makin kurang akurat.
4. *Area under the curve* adalah ukuran akurasi tes. Dia bervariasi antara 1 (excellent) sampai 0,5 artinya gagal berfungsi sebagai alat tes diagnostik.

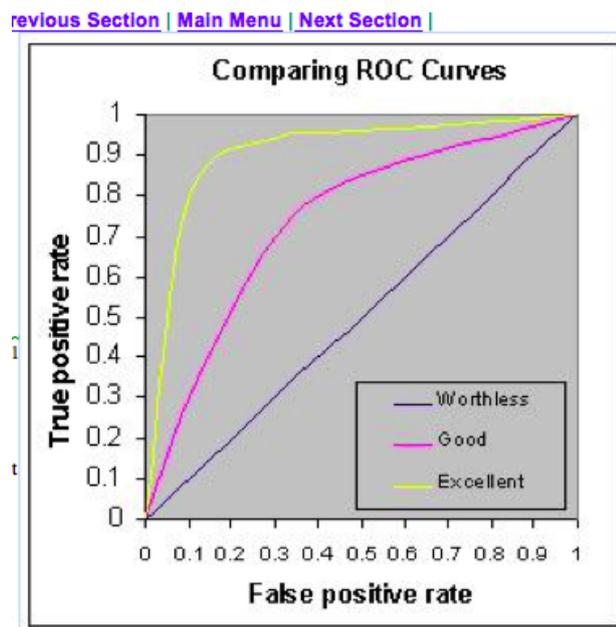
.90-1 = excellent

.80-.90 = good

.70-.80 = fair

.60-.70 = poor

.50-.60 = fail



Perbandingan 3 grafik ROC *curve*. Perhatikan bahwa garis paling mendekati tepi kiri dan atas adalah yang paling akurat, karena memberikan AUC paling besar. Garis yang

paling mendekati garis diagonal adalah yang paling tidak akurat karena memberikan AUC yang paling kecil (0.5).

1. Akobeng AK. Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta Paediatr.* 2007;96(3):338–41.